



PCT

**RAW SEQUENCE LISTING** DATE: 07/17/2003  
 PATENT APPLICATION: US/09/763,822A TIME: 11:32:33

Input Set : A:\W0008463.txt  
 Output Set: N:\CRF4\07172003\I763822A.raw

```

3 <110> APPLICANT: WALLAART, Thorvald Eelco
4      BOUWMEESTER, Hendrik Jan
6 <120> TITLE OF INVENTION: Transgenic Amorpha-4, 11-Diene Synthesis
8 <130> FILE REFERENCE: 702 010272
10 <140> CURRENT APPLICATION NUMBER: 09/763,822A
C--> 11 <141> CURRENT FILING DATE: 2003-05-05
13 <150> PRIOR APPLICATION NUMBER: PCT/EP99/06302
14 <151> PRIOR FILING DATE: 1999-08-27
16 <160> NUMBER OF SEQ ID NOS: 14
18 <170> SOFTWARE: MS Word 97 SR-2
20 <210> SEQ ID NO: 1
21 <211> LENGTH: 15
22 <212> TYPE: DNA
23 <213> ORGANISM: Artificial Sequence
25 <220> FEATURE:
26 <223> OTHER INFORMATION: EcoR I (Not I) adapter
28 <400> SEQUENCE: 1
29 gtcgacgcgg cccgcg 15
31 <210> SEQ ID NO: 2
32 <211> LENGTH: 19
33 <212> TYPE: DNA
34 <213> ORGANISM: Artificial Sequence
36 <220> FEATURE:
37 <223> OTHER INFORMATION: EcoR I (Not I) adapter
39 <400> SEQUENCE: 2
40 cagctgcgcc ggcgcattaa 19
42 <210> SEQ ID NO: 3
43 <211> LENGTH: 27
44 <212> TYPE: DNA
45 <213> ORGANISM: Artificial Sequence
47 <220> FEATURE:
48 <223> OTHER INFORMATION: Sense primer (primer C) used in PCR amplification
50 <400> SEQUENCE: 3
51 gtcgacaaaac catggcactt acagaag 27
53 <210> SEQ ID NO: 4
54 <211> LENGTH: 32
55 <212> TYPE: DNA
56 <213> ORGANISM: Artificial Sequence
58 <220> FEATURE:
59 <223> OTHER INFORMATION: Antisense primer (primer D) used in PCR amplification
61 <400> SEQUENCE: 4
62 ggatggatcc tcatatactc ataggataaa cg 32
64 <210> SEQ ID NO: 5

```

**ENTERED**

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/763,822A

DATE: 07/17/2003

TIME: 11:32:33

Input Set : A:\W0008463.txt

Output Set: N:\CRF4\07172003\I763822A.raw

65 <211> LENGTH: 23  
 66 <212> TYPE: DNA  
 67 <213> ORGANISM: Artificial Sequence  
 69 <220> FEATURE:  
 70 <223> OTHER INFORMATION: Sense primer (primer G) used in PCR amplification  
 72 <400> SEQUENCE: 5  
 73 gaggatccat gtcacttaca gaa 23  
 75 <210> SEQ ID NO: 6  
 76 <211> LENGTH: 24  
 77 <212> TYPE: DNA  
 78 <213> ORGANISM: Artificial Sequence  
 80 <220> FEATURE:  
 81 <223> OTHER INFORMATION: Antisense primer (primer H) used in PCR amplification  
 83 <400> SEQUENCE: 6  
 84 atggatcc tcatactcat agga 24  
 86 <210> SEQ ID NO: 7  
 87 <211> LENGTH: 22  
 88 <212> TYPE: DNA  
 89 <213> ORGANISM: Artificial Sequence  
 91 <220> FEATURE:  
 92 <223> OTHER INFORMATION: Sense primer (primer E) used in PCR amplification  
 94 <400> SEQUENCE: 7  
 95 cgagaatca tgtca ttac ag 22  
 97 <210> SEQ ID NO: 8  
 98 <211> LENGTH: 22  
 99 <212> TYPE: DNA  
 100 <213> ORGANISM: Artificial Sequence  
 102 <220> FEATURE:  
 103 <223> OTHER INFORMATION: Antisense primer (primer F) used in PCR amplification  
 105 <400> SEQUENCE: 8  
 106 ggatctcgag tcataatactc at 22  
 108 <210> SEQ ID NO: 9  
 109 <211> LENGTH: 538  
 110 <212> TYPE: DNA  
 111 <213> ORGANISM: Artificial Sequence  
 113 <220> FEATURE:  
 114 <223> OTHER INFORMATION: Nucleotide sequence of probe generated by PCR with primers A  
 115 and B  
 117 <400> SEQUENCE: 9  
 118 gatgagaatg gaaatttaa ggaatcgta gcta atgatg ttgaagg ttt gcttgagg 60  
 120 tacgaagcaa ctcttatgag ggtacctggg gagattat tagaagatgc tcttgg 120  
 122 acacgatctc gtcttagcat tatgacaaaaa gatgctttt ctacaaaccc cgctt 180  
 124 accgaaatac aacgggact aaagcaaccc cttggaaaaa gggtgccaag aata gaggcg 240  
 126 ggcgactaca ttcccttcta tcaacaacaa gattctcata acaagacttt actt 300  
 128 gctaagttag agttcaattt gcttcagtca ttgcacaagg aagagctcag ccatgtgtgc 360  
 130 aaatggtgg aagctttcga tatcaagaag aacgcaccc tttaagaga tagaatttt 420  
 132 gaatgtact ttggggact aggttcaggc tatgagccac agtattcccg ggcttaggtt 480  
 134 ttcttcacaa aagctgttgc tggataact cttatagacg acaccc tcga cgctacgg 538  
 136 <210> SEQ ID NO: 10

**RAW SEQUENCE LISTING**  
**PATENT APPLICATION: US/09/763,822A**

**DATE: 07/17/2003**  
**TIME: 11:32:33**

**Input Set : A:\W0008463.txt**  
**Output Set: N:\CRF4\07172003\I763822A.raw**

137 <211> LENGTH: 179  
 138 <212> TYPE: PRT  
 139 <213> ORGANISM: Artificial Sequence  
 141 <220> FEATURE:  
 142 <223> OTHER INFORMATION: Deduced amino acid sequence of probe generated by PCR with  
 143 primers A and B  
 145 <400> SEQUENCE: 10  
 146 Asp Glu Asn Gly Lys Phe Lys Glu Ser Leu Ala Asn Asp Val Glu Gly  
 147 1 5 10 15  
 148 Leu Leu Glu Leu Tyr Glu Ala Thr Ser Met Arg Val Pro Gly Glu Ile  
 149 20 25 30  
 150 Ile Leu Glu Asp Ala Leu Gly Phe Thr Arg Ser Arg Leu Ser Ile Met  
 151 35 40 45  
 152 Thr Lys Asp Ala Phe Ser Thr Asn Pro Ala Leu Phe Thr Glu Ile Gln  
 153 50 55 60  
 154 Arg Ala Leu Lys Gln Pro Leu Trp Lys Arg Leu Pro Arg Ile Glu Ala  
 155 65 70 75 80  
 156 Ala Gln Tyr Ile Pro Phe Tyr Gln Gln Asp Ser His Asn Lys Thr  
 157 85 90 95  
 158 Leu Leu Lys Leu Ala Lys Leu Glu Phe Asn Leu Leu Gln Ser Leu His  
 159 100 105 110  
 160 Lys Glu Glu Leu Ser His Val Cys Lys Trp Trp Lys Ala Phe Asp Ile  
 161 115 120 125  
 162 Lys Lys Asn Ala Pro Cys Leu Arg Asp Arg Ile Val Glu Cys Tyr Phe  
 163 130 135 140  
 164 Trp Gly Leu Gly Ser Gly Tyr Glu Pro Gln Tyr Ser Arg Ala Arg Val  
 165 145 150 155 160  
 166 Phe Phe Thr Lys Ala Val Ala Val Ile Thr Leu Ile Asp Asp Thr Phe  
 167 165 170 175  
 168 Asp Ala Thr  
 170 <210> SEQ ID NO: 11  
 171 <211> LENGTH: 2112  
 172 <212> TYPE: DNA  
 173 <213> ORGANISM: Artemisia annua L.  
 175 <220> FEATURE:  
 176 <223> OTHER INFORMATION: Nucleotide sequence of a positive clone (amorphadiene synthase  
 177 encoding gene) isolated from the cDNA library of induced A.annua  
 179 <400> SEQUENCE: 11  
 180 aattcgcggc cgcgctcgaca aatcatgtca cttacagaag aaaaacctat tcgccccatt 60  
 182 gccaactttc ctccaaggcat ttggggagat cagtttctca tctataaaa gcaagttagag 120  
 184 caagggttgg aacagatagt gaatgattta aaaaaagaag tgccgcaact actaaaagaa 180  
 186 gctttggata ttccatgaa acatgccaat ttgttgaagc tgattgtat aattcaacgc 240  
 188 cttgaaatac cgtatcactt tgaacggag attgatcatg cattgcaatg tatttatgaa 300  
 190 acatatggtg ataactggaa tggtgaccgc tcttcctt ggttccgtct tatgcgaaag 360  
 192 caaggatatt atgttacatg tcatgtttc aataactata aagacaaaaa tggagcggtc 420  
 194 aagcaatcgtagtcaatga tggtaaggt ttgttgcgt tgcgtacgc aacttctatg 480  
 196 agggtacctg gggagattat attagaagat gctttgggtt ttacacgatc tcgtctttagc 540  
 198 attatgacaa aagatgcttt ttctacaac cccgctctt ttaccgaaat acaacgggca 600  
 200 ctaaagcaac ccctttggaa aaggttgcca agaatagagg cggcgcagta cattccttgc 660

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/763,822A

DATE: 07/17/2003

TIME: 11:32:33

Input Set : A:\W0008463.txt

Output Set: N:\CRF4\07172003\I763822A.raw

**RAW SEQUENCE LISTING**  
**PATENT APPLICATION: US/09/763,822A**

**DATE: 07/17/2003**  
**TIME: 11:32:33**

**Input Set : A:\W0008463.txt**  
**Output Set: N:\CRF4\07172003\I763822A.raw**

278	115	120	125
279	Val Phe Asn Asn Tyr Lys Asp Lys Asn Gly Ala Phe Lys Gln Ser Leu		
280	130	135	140
281	Ala Asn Asp Val Glu Gly Leu Leu Glu Leu Tyr Glu Ala Thr Ser Met		
282	145	150	155
283	Arg Val Pro Gly Glu Ile Ile Leu Glu Asp Ala Leu Gly Phe Thr Arg		160
284	165	170	175
285	Ser Arg Leu Ser Ile Met Thr Lys Asp Ala Phe Ser Thr Asn Pro Ala		
286	180	185	190
287	Leu Phe Thr Glu Ile Gln Arg Ala Leu Lys Gln Pro Leu Trp Lys Arg		
288	195	200	205
289	Leu Pro Arg Ile Glu Ala Ala Gln Tyr Ile Pro Phe Tyr Gln Gln Gln		
290	210	215	220
291	Asp Ser His Asn Lys Thr Leu Leu Lys Leu Ala Lys Leu Glu Phe Asn		
292	225	230	235
293	Leu Leu Gln Ser Leu His Lys Glu Glu Leu Ser His Val Cys Lys Trp		240
294	245	250	255
295	Trp Lys Ala Phe Asp Ile Lys Lys Asn Ala Pro Cys Leu Arg Asp Arg		
296	260	265	270
297	Ile Val Glu Cys Tyr Phe Trp Gly Leu Gly Ser Gly Tyr Glu Pro Gln		
298	275	280	285
299	Tyr Ser Arg Ala Arg Val Phe Thr Lys Ala Val Ala Val Ile Thr		
300	290	295	300
301	Leu Ile Asp Asp Thr Tyr Asp Ala Tyr Gly Thr Tyr Glu Glu Leu Lys		
302	305	310	315
303	Ile Phe Thr Glu Ala Val Glu Arg Trp Ser Ile Thr Cys Leu Asp Thr		320
304	325	330	335
305	Leu Pro Glu Tyr Met Lys Pro Ile Tyr Lys Leu Phe Met Asp Thr Tyr		
306	340	345	350
307	Thr Glu Met Glu Glu Phe Leu Ala Lys Glu Gly Arg Thr Asp Leu Phe		
308	355	360	365
309	Asn Cys Gly Lys Glu Phe Val Lys Glu Phe Val Arg Asn Leu Met Val		
310	370	375	380
311	Glu Ala Lys Trp Ala Asn Glu Gly His Ile Pro Thr Thr Glu Glu His		
312	385	390	395
313	Asp Pro Val Val Ile Ile Thr Gly Gly Ala Asn Leu Leu Thr Thr Thr		400
314	405	410	415
315	Cys Tyr Leu Gly Met Ser Asp Ile Phe Thr Lys Glu Ser Val Glu Trp		
316	420	425	430
317	Ala Val Ser Ala Pro Pro Leu Phe Arg Tyr Ser Gly Ile Leu Gly Arg		
318	435	440	445
319	Arg Leu Asn Asp Leu Met Thr His Lys Ala Glu Gln Glu Arg Lys His		
320	450	455	460
321	Ser Ser Ser Ser Leu Glu Ser Tyr Met Lys Glu Tyr Asn Val Asn Glu		
322	465	470	475
323	Glu Tyr Ala Gln Thr Leu Ile Tyr Lys Glu Val Glu Asp Val Trp Lys		480
324	485	490	495
325	Asp Ile Asn Arg Glu Tyr Leu Thr Thr Lys Asn Ile Pro Arg Pro Leu		
326	500	505	510

**VERIFICATION SUMMARY**

PATENT APPLICATION: US/09/763,822A

DATE: 07/17/2003

TIME: 11:32:34

Input Set : A:\W0008463.txt

Output Set: N:\CRF4\07172003\I763822A.raw

L:11 M:271 C: Current Filing Date differs, Replaced Current Filing Date

SEQUENCE LISTING

<110> WALLAART, Thorvald Eelco  
BOUWMEESTER, Hendrik Jan

<120> Transgenic Amorpha-4, 11-Diene Synthesis

<130> 702 010272

<140> 09/763,822  
<141> 2001-02-26

<150> PCT/EP99/06302  
<151> 1999-08-27

<160> 14

<170> MS Word 97 SR-2

<210> 1  
<211> 15  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> EcoR I (Not I) adapter

<400> 1  
gtcgacgcgg ccgcg 15

<210> 2  
<211> 19  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> EcoR I (Not I) adapter

<400> 2  
cagctgcgcc ggcgcttaa 19

<210> 3  
<211> 27  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Sense primer (primer C) used in PCR amplification

<400> 3  
gtcgacaaac catggcactt acagaag 27

<210> 4  
<211> 32  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Antisense primer (primer D) used in PCR amplification

<400> 4  
ggatggatcc tcataatactc ataggataaa cg 32

<210> 5  
<211> 23  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Sense primer (primer G) used in PCR amplification

<400> 5  
gaggatccat gtcacttaca gaa 23

<210> 6  
<211> 24  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Antisense primer (primer H) used in PCR amplification

<400> 6  
atggatcctc atataactcat agga 24

<210> 7  
<211> 22  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Sense primer (primer E) used in PCR amplification

<400> 7  
cgagaattca tgtcacttac ag 22

<210> 8  
<211> 22  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Antisense primer (primer F) used in PCR amplification

<400> 8  
ggatctcgag tcataatactc at 22

<210> 9  
<211> 538  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Nucleotide sequence of probe generated by PCR with primers A

and B

<400> 9  
gatgagaatg ggaaatttaa ggaatcgta gctaatgatg ttgaaggttt gcttgagttg 60  
tacgaagcaa cttctatgag ggtacctggg gagattatat tagaagatgc tcttggttt 120  
acacgatctc gtcttagcat tatgacaaaa gatgctttt ctacaaaccc cgcttttt 180  
accgaaatac aacgggcact aaagcaaccc ctttgaaaaa ggttgccaag aatagaggcg 240  
gcgcagtaca ttccttcta tcaacaacaa gattctcata acaagacttt acttaaactt 300  
gctaaatggtag agttcaattt gtttcagtc ttgcacaagg aagagctcag ccatgtgtgc 360  
aaatggtgg aagcttcga tatcaagaag aacgcacctt gtttaagaga tagaattgtt 420  
gaatgctact tttggggact aggttcagc tatgagccac agtattcccgg ggttagagtt 480  
ttcttcacaa aagctgttgc tggtataact cttatagacg acacccctcga cgctacgg 538

<210> 10  
<211> 179  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Deduced amino acid sequence of probe generated by PCR with  
primers A and B

<400> 10  
Asp Glu Asn Gly Lys Phe Lys Glu Ser Leu Ala Asn Asp Val Glu Gly  
1 5 10 15  
Leu Leu Glu Leu Tyr Glu Ala Thr Ser Met Arg Val Pro Gly Glu Ile  
20 25 30  
Ile Leu Glu Asp Ala Leu Gly Phe Thr Arg Ser Arg Leu Ser Ile Met  
35 40 45  
Thr Lys Asp Ala Phe Ser Thr Asn Pro Ala Leu Phe Thr Glu Ile Gln  
50 55 60  
Arg Ala Leu Lys Gln Pro Leu Trp Lys Arg Leu Pro Arg Ile Glu Ala  
65 70 75 80  
Ala Gln Tyr Ile Pro Phe Tyr Gln Gln Asp Ser His Asn Lys Thr  
85 90 95  
Leu Leu Lys Leu Ala Lys Leu Glu Phe Asn Leu Leu Gln Ser Leu His  
100 105 110  
Lys Glu Glu Leu Ser His Val Cys Lys Trp Trp Lys Ala Phe Asp Ile  
115 120 125  
Lys Lys Asn Ala Pro Cys Leu Arg Asp Arg Ile Val Glu Cys Tyr Phe  
130 135 140  
Trp Gly Leu Gly Ser Gly Tyr Glu Pro Gln Tyr Ser Arg Ala Arg Val  
145 150 155 160  
Phe Phe Thr Lys Ala Val Ala Val Ile Thr Leu Ile Asp Asp Thr Phe  
165 170 175  
Asp Ala Thr

<210> 11  
<211> 2112

<212> DNA  
<213> Artemisia annua L.

<220>  
<223> Nucleotide sequence of a positive clone (amorphadiene synthase encoding gene) isolated from the cDNA library of induced A.annua

<400> 11  
aattcgcggc cgcgtcgaca aatcatgtca cttacagaag aaaaacctat tcgccccatt 60  
gccaacttgc tcccaaggcat ttggggagat cagttctca tctatcaaaa gcaagttagag 120  
caagggttgg aacagatagt gaatgattta aaaaaagaag tgccggcaact actaaaagaa 180  
gctttggata ttccatgaa acatgccaat ttgttgaagc tgattgtga aattcaacgc 240  
cttggaaatac cgtatcactt tgaacgggag attgatcatg cattgcaatg tatttatgaa 300  
acatatggtg ataactggaa tggtgaccgc tcttccttat ggttccgtct tatgcgaaag 360  
caaggatatt atgttacatg tggatgtttc aataactata aagacaaaaa tggagcgttc 420  
aagcaatcgt tagctaata tggatgttgc ttttccttat ggttccgtct tatgcgaaag 480  
agggtacctg gggagattat attagaagat gctcttggtt ttacacgatc tcgtcttagc 540  
attatgacaa aagatgcttt ttctacaaac cccgctcttt ttaccgaaat acaacgggca 600  
ctaaagcaac ccctttggaa aaggttgcca agaatagagg cggcgcagta cattccttgc 660  
tatcaacaac aagattctca taacaagact ttacttaaac ttgtctaagtt agagttaat 720  
ttgcttcagt cattgcacaa ggaagagctc agccatgtgt gcaaattggtg gaaagcttgc 780  
gatatacaaga agaacgcacc ttgtttaaga gatagaattt ttgaatgcta ctgttgggaa 840  
ctaggttcag gctatgagcc acagtattcc cgggcttagag ttttcttcac aaaagctgtt 900  
gctgttataa ctcttataga tgacacttat gatgcgtatg gtacttatga agaacttaag 960  
atctttactg aagctgttga aagggttgtca attacatgtc tagacacact tccagaatac 1020  
atgaaaccga tatacaaattt attcatggat acatacacag aaatggaaat atttcttgca 1080  
aaggagggaa gaacagatct atttaactgc ggcaaaagaat ttgtgaaaga gtttgttaga 1140  
aacctgatgg ttgaagcaaa atgggcaaat gagggacaca taccaaccac tgaagagcat 1200  
gatccagttg taatcattac tggcggtgt aacctgctta caacaacttg ttatcttggc 1260  
atgagtgata tattcacaaa agagtctgtc gaatgggctg tctctgcacc tcctctttt 1320  
agataactcag gtataacttgg tcgacgccta aatgatctca tgacccacaa ggccgagcaa 1380  
gaaagaaaaac atagttcatc gagccttgaa agtttatatga aggaatataa tgtcaatgag 1440  
gagttatgccca aaaccttgat ttacaaggaa gtagaaagatg tggaaagaataa tataaaccga 1500

gagttaccta caactaaaaa cattccaagg ccgttattga tggctgtat ctatttgtc 1560  
cagtttcttg aagttcaata tgccggaaag gataacttca cacgtatggg agacgaatac 1620  
aaacatctca taaagtctct actcgtttat cctatgagta tatgactacc aatccttcgt 1680  
gcatacgctta tcaattatat taaaagggtt aactatgcac gtctctatgg agagaatttc 1740  
tcaagctatt tggtgttct tgcggcaat aataaatcag acgcataaaa ttgtattgaa 1800  
ctatatgccg atagctatTTT aaagttatta tacaactaaa atattcaaca atggattat 1860  
actttactt tgtacaaaag caaaagtaca ctactgttat gtaacatTTT agttctatga 1920  
tacTTtagtt acgaatcggc ttatatacat tgatacactt ttatgcagaa aaccctagta 1980  
aataaaaagt cgatatcttg tactacacat atcgacgaa tttccgtttc ccgtttgtat 2040  
tttacgatAT gttatttaat gaatatgttt catgtggttg ttgcttaaaa aaaaagtgcg 2100  
cgccggccgcg aa 2112

<210> 12

<211> 697

<212> PRT

<213> Artemisia annua L.

<220>

<223> Deduced amino acid sequence of a positive clone (amorphadiene synthase encoding gene) isolated from the cDNA library of induced A.annua

<400> 12

Asn	Ser	Arg	Pro	Arg	Arg	Gln	Ile	Met	Ser	Leu	Thr	Glu	Glu	Lys	Pro
1						5					10			15	
Ile	Arg	Pro	Ile	Ala	Asn	Phe	Pro	Pro	Ser	Ile	Trp	Gly	Asp	Gln	Phe
			20							25				30	
Leu	Ile	Tyr	Gln	Lys	Gln	Val	Glu	Gln	Gly	Val	Glu	Gln	Ile	Val	Asn
			35				40				45				
Asp	Leu	Lys	Lys	Glu	Val	Arg	Gln	Leu	Leu	Lys	Glu	Ala	Leu	Asp	Ile
	50					55				60					
Pro	Met	Lys	His	Ala	Asn	Leu	Leu	Lys	Leu	Ile	Asp	Glu	Ile	Gln	Arg
	65					70				75			80		
Leu	Gly	Ile	Pro	Tyr	His	Phe	Glu	Arg	Glu	Ile	Asp	His	Ala	Leu	Gln
					85				90			95			
Cys	Ile	Tyr	Glu	Thr	Tyr	Gly	Asp	Asn	Trp	Asn	Gly	Asp	Arg	Ser	Ser
						100			105			110			
Leu	Trp	Phe	Arg	Leu	Met	Arg	Lys	Gln	Gly	Tyr	Tyr	Val	Thr	Cys	Asp
		115					120				125				
Val	Phe	Asn	Asn	Tyr	Lys	Asp	Lys	Asn	Gly	Ala	Phe	Lys	Gln	Ser	Leu
		130				135			140						
Ala	Asn	Asp	Val	Glu	Gly	Leu	Leu	Glu	Leu	Tyr	Glu	Ala	Thr	Ser	Met
	145					150				155			160		
Arg	Val	Pro	Gly	Glu	Ile	Ile	Leu	Glu	Asp	Ala	Leu	Gly	Phe	Thr	Arg
					165				170			175			
Ser	Arg	Leu	Ser	Ile	Met	Thr	Lys	Asp	Ala	Phe	Ser	Thr	Asn	Pro	Ala

	180	185	190												
Leu	Phe	Thr	Glu	Ile	Gln	Arg	Ala	Leu	Lys	Gln	Pro	Leu	Trp	Lys	Arg
195								200				205			
Leu	Pro	Arg	Ile	Glu	Ala	Ala	Gln	Tyr	Ile	Pro	Phe	Tyr	Gln	Gln	Gln
210					215					220					
Asp	Ser	His	Asn	Lys	Thr	Leu	Leu	Lys	Leu	Ala	Lys	Leu	Glu	Phe	Asn
225					230					235			240		
Leu	Leu	Gln	Ser	Leu	His	Lys	Glu	Glu	Leu	Ser	His	Val	Cys	Lys	Trp
					245				250			255			
Trp	Lys	Ala	Phe	Asp	Ile	Lys	Lys	Asn	Ala	Pro	Cys	Leu	Arg	Asp	Arg
					260				265			270			
Ile	Val	Glu	Cys	Tyr	Phe	Trp	Gly	Leu	Gly	Ser	Gly	Tyr	Glu	Pro	Gln
					275			280			285				
Tyr	Ser	Arg	Ala	Arg	Val	Phe	Phe	Thr	Lys	Ala	Val	Ala	Val	Ile	Thr
					290			295			300				
Leu	Ile	Asp	Asp	Thr	Tyr	Asp	Ala	Tyr	Gly	Thr	Tyr	Glu	Glu	Leu	Lys
305						310				315			320		
Ile	Phe	Thr	Glu	Ala	Val	Glu	Arg	Trp	Ser	Ile	Thr	Cys	Leu	Asp	Thr
					325				330			335			
Leu	Pro	Glu	Tyr	Met	Lys	Pro	Ile	Tyr	Lys	Leu	Phe	Met	Asp	Thr	Tyr
					340				345			350			
Thr	Glu	Met	Glu	Glu	Phe	Leu	Ala	Lys	Glu	Gly	Arg	Thr	Asp	Leu	Phe
					355			360			365				
Asn	Cys	Gly	Lys	Glu	Phe	Val	Lys	Glu	Phe	Val	Arg	Asn	Leu	Met	Val
					370			375			380				
Glu	Ala	Lys	Trp	Ala	Asn	Glu	Gly	His	Ile	Pro	Thr	Thr	Glu	Glu	His
385						390				395			400		
Asp	Pro	Val	Val	Ile	Ile	Thr	Gly	Gly	Ala	Asn	Leu	Leu	Thr	Thr	Thr
					405				410			415			
Cys	Tyr	Leu	Gly	Met	Ser	Asp	Ile	Phe	Thr	Lys	Glu	Ser	Val	Glu	Trp
					420				425			430			
Ala	Val	Ser	Ala	Pro	Pro	Leu	Phe	Arg	Tyr	Ser	Gly	Ile	Leu	Gly	Arg
					435				440			445			
Arg	Leu	Asn	Asp	Leu	Met	Thr	His	Lys	Ala	Glu	Gln	Glu	Arg	Lys	His
					450			455			460				
Ser	Ser	Ser	Ser	Leu	Glu	Ser	Tyr	Met	Lys	Glu	Tyr	Asn	Val	Asn	Glu
465						470				475			480		
Glu	Tyr	Ala	Gln	Thr	Leu	Ile	Tyr	Lys	Glu	Val	Glu	Asp	Val	Trp	Lys
					485				490			495			
Asp	Ile	Asn	Arg	Glu	Tyr	Leu	Thr	Thr	Lys	Asn	Ile	Pro	Arg	Pro	Leu
					500				505			510			
Leu	Met	Ala	Val	Ile	Tyr	Leu	Cys	Gln	Phe	Leu	Glu	Val	Gln	Tyr	Ala
					515			520			525				
Gly	Lys	Asp	Asn	Phe	Thr	Arg	Met	Gly	Asp	Glu	Tyr	Lys	His	Leu	Ile
					530			535			540				
Lys	Ser	Leu	Leu	Val	Tyr	Pro	Met	Ser	Ile	Leu	Pro	Ile	Leu	Arg	Ala
545						550				555			560		
Pro	Ile	Asn	Tyr	Ile	Glu	Arg	Val	Asn	Tyr	Ala	Arg	Leu	Tyr	Gly	Glu
					565				570			575			
Asn	Phe	Ser	Ser	Tyr	Leu	Val	Phe	Leu	Ala	Gly	Asn	Asn	Lys	Ser	Asp
					580				585			590			
Ala	Asn	Cys	Ile	Glu	Leu	Tyr	Ala	Asp	Ser	Tyr	Leu	Lys	Leu	Leu	Tyr
					595				600			605			
Asn	Asn	Ile	Gln	Gln	Trp	Tyr	Tyr	Thr	Phe	Thr	Leu	Tyr	Lys	Ser	Lys
					610			615			620				
Ser	Thr	Leu	Leu	Leu	Cys	Asn	Ile	Leu	Val	Leu	Tyr	Phe	Ser	Tyr	Glu
					625			630			635			640	

Ser Ala Tyr Ile His Tyr Thr Phe Met Gln Lys Thr Leu Val Asn Lys  
645 650 655  
Lys Ser Ile Ser Cys Thr Thr His Ile Ala Arg Ile Ser Val Cys g  
660 665 670  
Leu Tyr Phe Thr Ile Cys Tyr Leu Met Asn Met Phe His Val Val Val  
675 680 685  
Ala Lys Lys Ser Arg Arg Gly Arg Glu  
690 695

<210> 13

<211> 1649

<212> DNA

<213> Artificial Sequence

<220>

<223> Nucleotide sequence of the amorphadiene synthase encoding gene,  
between start and stop codon, obtained by PCR with primers C  
and D

<400> 13

ccatggcact tacagaagaa aaacctattc gccccattgc caacttcct ccaagcattt 60  
ggggagatca gtttctcatc tatcaaaagc aagtagagca aggggtggaa cagatagtga 120  
atgatttaaa aaaagaagtg cgccaactac taaaagaagc ttggatatt cctatgaaac 180  
atgccaattt gttgaagctg attgatgaaa ttcaacgcct tggaaataccg tatcactttg 240  
aacgggagat tgatcatgca ttgcaatgta tttatgaaac atatggat aactggaatg 300  
gtgaccgctc ttccattatgg ttccgtctta tgcgaaagca aggatattat gttacatgtg 360  
atgtttcaa taactataaa gacaaaaatg gagcgttcaa gcaatcgta gctaattatg 420  
ttgaaggttt gcttgagttt tacgaagcaa cttctatgag ggtacctggg gagattat 480  
tagaagatgc tcttggttt acacgatctc gtcttagcat tatgacaaaa gatgttttt 540  
ctacaaaccc cgctctttt accgaaatac aacggcact aaagcaaccc ctggaaaa 600  
ggttgccaag aatagaggcg ggcgactaca ttccattctta tcaacaacaa gattctcata 660  
acaagacttt acttaaactt gctaagttt agttcaattt gtttgcgtca ttgcacaagg 720  
aagagctcg ccatgtgtgc aaatggtgg aagcttcga tatcaagaag aacgcacctt 780  
gtttaagaga tagaattgtt gaatgtact tttggggact aggttcaggc tatgagccac 840  
agtattcccg ggcttagagtt ttcttcacaa aagctgtgc tttataact cttatagatg 900  
acacttatga tgcgtatggt acttatgaag aacttaagat ctttactgaa gctgtgaaa 960  
ggtgtcaat tacatgctt gacacacttc cagaatacat gaaaccgata tacaattat 1020  
tcatggatac atacacagaa atgaaagaat ttcttgcaaa ggaggaaaga acagatctat 1080  
ttaactgcgg caaagaattt gtgaaagagt ttgttagaaa cctgtatgggtaa gaagcaaaat 1140

ggccaaatga gggacacata ccaaccactg aagagcatga tccagttgt atcattactg 1200  
gcggtgctaa cctgcttaca acaacttgtt atcttggcat gagtgatata ttcacaaaag 1260  
agtctgtcga atgggctgtc tctgcaccc tccttttag atactcaggat atactggtc 1320  
gacgcctaaa tgatctcatg acccacaagg ccgagcaaga aagaaaacat agttcatcga 1380  
gccttggaaag ttatataatg gaatataatg tcaatgagga gtatgccaa accttgattt 1440  
acaaggaagt agaagatgtg tgaaagata taaaccgaga gtacctcaca actaaaaaca 1500  
ttccaaggcc gttattgatg gctgtgatct atttgcgcgaa gttcaatatg 1560  
cagggaaagga taacttcaca cgtatggag acgaatacaa acatctcata aagtctctac 1620  
tcgttttatcc tatgagtgata tgaggatcc 1649

<210> 14

<211> 549

<212> PRT

<213> Artificial Sequence

<220>

<223> Deduced amino acid sequence of the amorphadiene synthase encoding gene, between start and stop codon, obtained by PCR with primers C and D

<400> 14

Thr	Met	Ala	Leu	Thr	Glu	Glu	Lys	Pro	Ile	Arg	Pro	Ile	Ala	Asn	Phe
					5				10					15	
Pro	Pro	Ser	Ile	Trp	Gly	Asp	Gln	Phe	Leu	Ile	Tyr	Gln	Lys	Gln	Val
					20				25				30		
Glu	Gln	Gly	Val	Glu	Gln	Ile	Val	Asn	Asp	Leu	Lys	Lys	Glu	Val	Arg
					35			40			45				
Gln	Leu	Leu	Lys	Glu	Ala	Leu	Asp	Ile	Pro	Met	Lys	His	Ala	Asn	Leu
					50			55			60				
Leu	Lys	Leu	Ile	Asp	Glu	Ile	Gln	Arg	Leu	Gly	Ile	Pro	Tyr	His	Phe
					65		70			75			80		
Glu	Arg	Glu	Ile	Asp	His	Ala	Leu	Gln	Cys	Ile	Tyr	Glu	Thr	Tyr	Gly
					85			90			95				
Asp	Asn	Trp	Asn	Gly	Asp	Arg	Ser	Ser	Leu	Trp	Phe	Arg	Leu	Met	Arg
					100			105			110				
Lys	Gln	Gly	Tyr	Tyr	Val	Thr	Cys	Asp	Val	Phe	Asn	Asn	Tyr	Lys	Asp
					115			120			125				
Lys	Asn	Gly	Ala	Phe	Lys	Gln	Ser	Leu	Ala	Asn	Asp	Val	Glu	Gly	Leu
					130			135			140				
Leu	Glu	Leu	Tyr	Glu	Ala	Thr	Ser	Met	Arg	Val	Pro	Gly	Glu	Ile	Ile
					145			150			155			160	
Leu	Glu	Asp	Ala	Leu	Gly	Phe	Thr	Arg	Ser	Arg	Leu	Ser	Ile	Met	Thr
					165			170			175				
Lys	Asp	Ala	Phe	Ser	Thr	Asn	Pro	Ala	Leu	Phe	Thr	Glu	Ile	Gln	Arg
					180			185			190				
Ala	Leu	Lys	Gln	Pro	Leu	Trp	Lys	Arg	Leu	Pro	Arg	Ile	Glu	Ala	Ala
					195			200			205				
Gln	Tyr	Ile	Pro	Phe	Tyr	Gln	Gln	Gln	Asp	Ser	His	Asn	Lys	Thr	Leu

210	215	220
Leu Lys Leu Ala Lys Leu Glu Phe Asn Leu Leu Gln Ser Leu His Lys		
225	230	235
Glu Glu Leu Ser His Val Cys Lys Trp Trp Lys Ala Phe Asp Ile Lys		240
245	250	255
Lys Asn Ala Pro Cys Leu Arg Asp Arg Ile Val Glu Cys Tyr Phe Trp		
260	265	270
Gly Leu Gly Ser Gly Tyr Glu Pro Gln Tyr Ser Arg Ala Arg Val Phe		
275	280	285
Phe Thr Lys Ala Val Ala Val Ile Thr Leu Ile Asp Asp Thr Tyr Asp		
290	295	300
Ala Tyr Gly Thr Tyr Glu Glu Leu Lys Ile Phe Thr Glu Ala Val Glu		
305	310	315
Arg Trp Ser Ile Thr Cys Leu Asp Thr Leu Pro Glu Tyr Met Lys Pro		
325	330	335
Ile Tyr Lys Leu Phe Met Asp Thr Tyr Thr Glu Met Glu Glu Phe Leu		
340	345	350
Ala Lys Glu Gly Arg Thr Asp Leu Phe Asn Cys Gly Lys Glu Phe Val		
355	360	365
Lys Glu Phe Val Arg Asn Leu Met Val Glu Ala Lys Trp Ala Asn Glu		
370	375	380
Gly His Ile Pro Thr Thr Glu Glu His Asp Pro Val Val Ile Ile Thr		
385	390	395
400		400
Gly Gly Ala Asn Leu Leu Thr Thr Cys Tyr Leu Gly Met Ser Asp		
405	410	415
Ile Phe Thr Lys Glu Ser Val Glu Trp Ala Val Ser Ala Pro Pro Leu		
420	425	430
Phe Arg Tyr Ser Gly Ile Leu Gly Arg Arg Leu Asn Asp Leu Met Thr		
435	440	445
His Lys Ala Glu Gln Glu Arg Lys His Ser Ser Ser Ser Leu Glu Ser		
450	455	460
Tyr Met Lys Glu Tyr Asn Val Asn Glu Glu Tyr Ala Gln Thr Leu Ile		
465	470	475
480		
Tyr Lys Glu Val Glu Asp Val Trp Lys Asp Ile Asn Arg Glu Tyr Leu		
485	490	495
Thr Thr Lys Asn Ile Pro Arg Pro Leu Leu Met Ala Val Ile Tyr Leu		
500	505	510
Cys Gln Phe Leu Glu Val Gln Tyr Ala Gly Lys Asp Asn Phe Thr Arg		
515	520	525
Met Gly Asp Glu Tyr Lys His Leu Ile Lys Ser Leu Leu Val Tyr Pro		
530	535	540
Met Ser Ile Gly Ser		
545		